

DOCUMENT RESUME

ED 172 789

IR 007 422

AUTHOR Roberts, Dennis M.
TITLE A Consumer's Guide to Common Flaws in Research.
PUB DATE Mar 79
NOTE 9p.; Paper presented at the Annual Convention of the Association for Educational Communications and Technology (New Orleans, Louisiana, March 1979); For related document, see IR 007 416

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Educational Research; *Research Design; *Research Methodology; *Research Problems; *Research Skills; Research Utilization

ABSTRACT

This brief paper presents a sequential walk through a simple research setting--a two group experiment--to indicate common but avoidable design difficulties that may jeopardize the usefulness of the data obtained. The six flaws examined are the failure to select subjects from the population to which generalizations are to be made; failure to use randomization in assigning students to experimental and control groups; failure to use an adequate number of students; failure to maintain control over manipulation of the independent variable; failure to use reliable and/or valid measures of the dependent variable; and failure to use an appropriate statistical test. Several additional flaws that may creep into research studies are indicated in the closing comments, and three references are listed. (Author/RAO)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

A Consumer's Guide to Common Flaws in Research *

Dennis M. Roberts

The Pennsylvania State University

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Michael Simonson

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
USERS OF THE ERIC SYSTEM."

* The present paper is a revised and condensed version
of a paper "Common Flaws In Research Design" by
William Rabinowitz and Dennis Roberts published
in the NABTE Journal, 1977, #4.

ED172789

IR007422

Planning the Research Study

The purpose of research is to shed light on important questions phrased in testable forms. In general, one seeks to establish functional or causal relationships between independent and dependent variables. However, designing studies to establish these relationships or causes involves many potential hazards. In any case though, the investigator has an obligation to plan the study in order to maximize the interpretability of the findings. Unfortunately, there are many basic flaws that can seriously hinder results interpretation. The present brief paper is a sequential walk through a simple research setting-- a two group experiment-- to indicate common (but avoidable) design difficulties that will jeopardize the usefulness of the obtained data.

Common Flaws

Figure 1 presents a flow chart type diagram of several important steps in the research process. Obviously, it is kept simple-- many details of actual design execution have been omitted because of space limitations. Assume, for

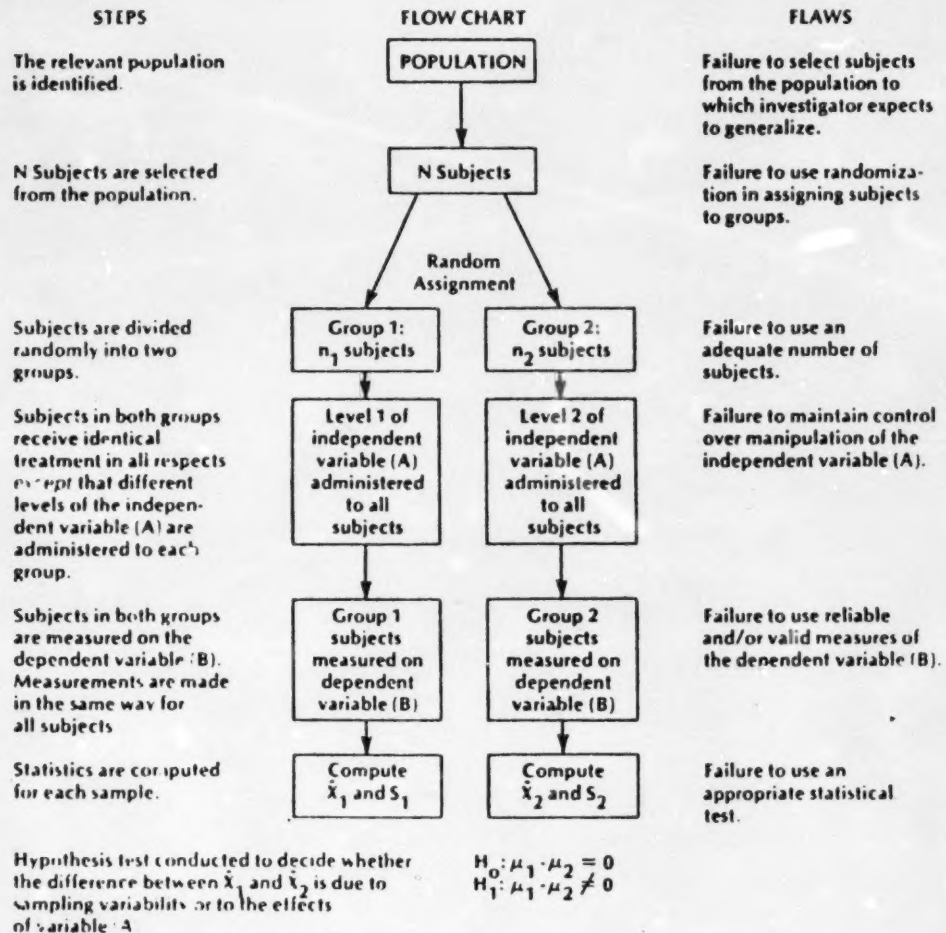
Insert Figure 1 about here

purposes of the remaining discussion, that "the" important question is "Will the use of electronic calculators improve elementary school students' attitudes towards mathematics?"

1. Failure to select subjects from the population to which generalizations are to be made.

This item deals specifically with what Campbell and Stanley (1963) refer to as external validity. Clearly, the question implies that an answer is

Figure 1. Steps and possible design flaws in an experiment to study the effect of treatment variable A on variable B.



desired for the overall population of elementary school students. Therefore, there should be evidence in the methods section that shows students at varying elementary grade levels are included in the sample. Secondly, there must be evidence that the students from these several grade levels were drawn (for the study) in a way to assure a representative sample. If the sample differs in important ways from the population, the results may not generalize to the population of interest as stated in the research question.

2. Failure to use randomization in assigning students to experimental and control groups.

It is absolutely essential that both groups be equivalent at the onset of the experiment. Differences between the groups on the dependent variable that appear at the conclusion of the study will invariably be interpreted as resulting from the impact of the independent variable. However, this conclusion cannot be justified if groups were different initially in important (and unknown) ways. Unfortunately, there is no way to absolutely guarantee equivalence of both groups initially. However, random assignment does ensure equivalence within the limits of chance. Since random fluctuations (chance aberrations) are estimatable, random assignment offers a way to control the initial makeup of the two groups. If the method section does not give specific indications that students were assigned at random to the two groups, the entire set of results are very suspect.

3. Failure to use an adequate number of students.

If the hope is to support the notion that use of calculators will improve attitudes towards mathematics, then we certainly want to reject the null hypothesis (μ Attitude with Calculators = μ Attitude without Calculators) if our notion is actually correct. In this situation, rejecting the null in the

study when our position is the true state of affairs is a correct decision to have reached. This type of correct decision is called the power of a statistical test. Unfortunately, unless your independent variable is very potent, power is greatly influenced by sample size. Also, unfortunately, the ability to build a certain degree of power into a study-- while doable-- is fairly complicated. A paper by Hopkins (1974) may be helpful to the interested reader. In the present study, it is doubtful that calculators are going to work miracles on students' attitudes towards mathematics. Thus, a failure to use a "sufficient" sized sample is certainly doomed to a retention of the null hypothesis.

4. Failure to maintain control over manipulation of the independent variable.

A necessary assumption is that students in the two groups were, in fact, treated differently throughout the period of the study. In our context, this requires that no control students have access to calculators during the experiment. While this may be controllable during the math classes, control students might have access to and use calculators in other classes (science?) and/or outside of school. Such contamination is very likely to reduce the differences between the two groups on the dependent variable. What will look like an ineffective independent variable actually disguises the possibility that the calculator did have an effect; however, inadvertently in both groups.

5. Failure to use reliable and/or valid measures of the dependent variable.

The original question posed stated that the important dependent variable of interest is "attitudes toward mathematics". Assuming for a moment that such a construct exists, then at issue is whether the construct can be measured consistently. Without some data being presented to demonstrate reliability of the measure in this particular research setting, then it is unlikely that differences between the groups on the dependent measure will

occur. Assuming reasonable reliability, however, a more important issue is whether the dependent measure in the study is, in fact, measuring attitudes toward mathematics. Obviously, any comparison between the groups on a measure that isn't tapping the important construct of interest is useless and certainly no test of the hypothesis we might have generated from our research question (i.e., calculators improve attitudes towards mathematics).

6. Failure to use an appropriate statistical test.

One area where flaws shouldn't occur but do quite often is the inappropriate use of certain statistical tests. The emphasis here is not on using a test that is less powerful than others but rather on using a test incorrectly. A common example here is using analysis of covariance (ANCOVA) to make statistical adjustments for the fact that groups were not equivalent initially. What this means (in all probability) is that a statistical technique is supposed to correct for non-randomization of students to the experimental and control groups. Unfortunately, ANCOVA can't make non-random assignment random; no statistical test can correct for that difficulty. ANCOVA has two sets of assumptions that need to be met in order to be interpreted properly -- there are those for analysis of variance (ANOVA) and regression. If random assignment has been made and there is a reasonable correlation between a covariate (measured-- incidentally-- before the experiment begins) and the dependent variable, then ANCOVA may reduce the error term, hence, giving a better chance to reject the null hypothesis. The interested reader is referred to an article on ANCOVA by Kennedy (1977) for elaborations on this topic.

Closing Comments

The present focus has been on several common flaws that occur in even the simplest research design. Obviously, many other flaws may creep into research studies such that the output data are rendered suspect. First, many studies do not use an appropriate research design to handle the hypotheses being tested. Second, many research questions are formulated in such a way that no data will be able to support the hypotheses. And third, many investigations are conducted without proper consideration of the relevant variables that might have an influence on the dependent measures. All these additional problems plus the others briefly outlined above pose threats to both the internal and external validity of research investigations. Careful design and control, plus a good measure of common sense, are the best ways in empirical studies to insure meaningful data from one's research efforts.

References

- Campbell, D., and Stanley, J. Experimental and quasi-experimental designs for research. Handbook of Research on Teaching, N. L. Gage, editor. Skokie, Ill.: Rand McNally, 1963.
- Hopkins, K. D. Preventing the number-one misinterpretation of behavioral research, or how to increase statistical power. Journal of Special Education, Volume 8, 1974.
- Kennedy, J. J. The use and misuse of analysis of covariance. NABTE Review, 1977.